ON GRADUATION OF AGE-SPECIFIC FERTILITY RATES

Syed Zamir Hassan

December 1973

Abstract

In this paper an attempt is made to graduate the age-specific fertility schedules by fitting a linear function using the method of least squares. From the analysis of the two different fertility schedules it appears that this method reveals quite good results and gives a reasonably good fit. The method suggested here is simpler and more efficient than the most commonly used method of moments and it seems that we can safely use it for the graduation of the fertility schedules.

In the Mimeo Series of the Biometrics Unit, Cornell University.

ON GRADUATION OF AGE-SPECIFIC FERTILITY RATES

Syed Zamir Hassan

December 1973

Avery (1970) and Mitra (1970, 1967) tried to evaluate the effectiveness of some of the most commonly suggested graduation functions for the Net Maternity Function and Age-Specific Fertility Rates. Both of the authors used the method of moments for the fitting of the graduation functions. Fisher showed that in general the method of moments does not give efficient estimators of population parameters.

In this paper an attempt is made to fit a linear function by the method of least squares to the Age-Specific fertility schedules. It appears that the method reveals quite good results and gives a reasonably good fit.

Avery (1970), after fitting seven different graduation functions to 392 fertility schedules, observed that the Pearsonian type 1 model is the best of all other graduation functions. The examination of moments of Age-Specific fertility schedule, particularly β_1 and β_2, also gives a justification of using the Pearsonian type 1 model as a theoretical model to graduate the Age-Specific fertility schedules.

The theory of type 1 model is straight forward, but its application is somewhat difficult in view of a large number of parameters.

The Pearsonian type 1 model is given by the following equation:

$$f(x) = A\left(1 + \frac{x}{a_1}\right)^{m_1}\left(1 - \frac{x}{a_2}\right)^{m_2} \quad -a_1 \le x \le a_2 \quad .$$

If we substitute $\frac{a_2-x}{a_2-a_1} = t$ the distribution of t turns out to be Beta with parameters $m_1 + 1$ and $m_2 + 1$, i.e.

$$f(t) = \frac{1}{B(m_1+1,m_2+1)} t^{m_2}(1-t)^{m_1} \quad 0 \leq t \leq 1 \quad .$$

It is well known that even for moderately large values of m_1 and m_2, ξ follows approximately a normal distribution with mean μ_ξ and variance σ_ξ^2, where

$$\xi = \log_e \left(\frac{t}{1-t} \right)$$

$$\mu_\xi = \log_e \left(\frac{m_2 + 1/2}{m_1 + 1/2} \right)$$

and

$$\sigma_\xi^2 = (m_1+1)^{-1} + (m_2+1)^{-1} \quad .$$

It is therefore advisable to try this transformation for fitting instead of fitting a type 1 model having a large number of parameters. This procedure will also avoid the difficulty of having to perform approximate integration as will be needed in fitting type 1 model because probability integral tables for standard normal variates are available.

As is obvious, an ambitious analysis on the above lines would require the estimation of a_1 and a_2. For practical purposes f(x) can be regarded as zero for $x < \alpha$ and $x > \beta$ for some α and β (say 50 and 15), where $\beta - \alpha = a_1 + a_2 = 35$ is assumed to be the range of reproduction age.

Now taking the advantage of the above discussed transformation, the cumulative age specific fertility rate m(x) may be reduced to the realtive percentage fertility rates $F_m(x)$ where

$$F_m(x) = \frac{m_i(x)}{m_7(x)} \qquad 1 \leq i \leq 7$$

and corresponding to these $F_m(x)$ the standard normal deviates z may be read off from any normal integral tables. If, now, the transformed variable ξ is normal then this, when plotted against the normal deviate z, should yield points lying on a straight line, given by

$$z = a + b\xi \quad .$$

Now, using the method of least squares, the estimates of a and b may be obtained. These estimates can be used to calculate estimated normal deviates $(\hat{z})$ for any given value of the transformed variate (ξ). These normal deviates $(\hat{z})$ will give the fitted values of the relative percentage fertility rates $[\hat{F}_m(x)]$.

The application of this method is demonstrated in the following table I and table II. In table I a fertility schedule of Pakistan is used and table II shows the computations made on a fertility schedule of France.

From the study of tables I and II it seems that the fit, using this method, is reasonably good and this method of fitting, which is simpler and more efficient than the method of moments, may safely be used for the graduation of fertility schedules.

Table I: Analysis of fertility schedule of the population of Karachi (Pakistan) 1958. [Hashmi, table V.3]

Age group X	$F_m(x)$ (observed)	z	ξ	$\hat{z}$	$\hat{F}_m(x)$ (expected)	$\lvert F_m(x)-\hat{F}_m(x)\rvert$
15-19	0.09567	-1.31	2.71	-1.42*	0.07780	0.01787
20-24	0.31207	-0.49	1.79	-0.53	0.29806	0.01401
25-29	0.55429	0.14	0.92	0.18	0.57142	0.01713
30-34	0.74563	0.66	0.28	0.71	0.76114	0.01551
35-39	0.89749	1.27	-0.29	1.19	0.88298	0.01451
40-44	0.95672	1.71	-0.92	1.72	0.95728	0.00056
45-49	1.00000	∞	-1.79	2.44	0.99266	0.007340

*For the first age group the minimum reproductive age is assumed 13 years.

$$\xi = \frac{\text{Max age} - x_u}{x_u - \text{Min age}} = \frac{50 - x_u}{x_u - 15}$$

$\hat{z} = \hat{a} + \hat{b}\xi$ where $\hat{a} = 0.9537$ and $\hat{b} = -0.8306$

To test the goodness of fit the Kodmogrove-Simerncve test is applied. The working of the test is as follows [Messay, 1951].

$$H_o : F_m(x) = \hat{F}_m(x)$$

Critical region: Reject H_o if $D_n \geq .03747$

Test Statistic : $D_n = \text{l.u.b.} \ |F_m(x) - \hat{F}_m(x)|$
$= 0.01787$.

Since we find that the test statistic D_n (observed) is smaller than the D_n tabulated at the 5 percent level of significance, then we can safely assume that the differences between the observed $F_m(x)$ and expected $F_m(x)$ is not significantly large.

Table II: Analysis of fertility schedule of the population of France 1956. [Pressat, table 8.6a]

Age group x	$F_m(x)$ (observed)	z	ξ	$\hat{z}$	$\hat{F}_m(x)$ (expected)	$\|F_m(x)-\hat{F}_m(x)\|$
15-19	0.04160	-1.73	2.71	-1.69	0.05050	0.00890
20-24	0.33315	-0.43	1.79	-0.56	0.28774	0.04541
25-29	0.64884	0.39	0.92	0.45	0.67364	0.02480
30-34	0.84940	1.03	0.28	1.20	0.88493	0.03553
35-39	0.96472	1.81	-0.29	1.87	0.96926	0.00454
40-44	0.99721	2.78	-0.92	2.61	0.99547	0.00174
45-49	1.00000	∞	-1.79	3.63	0.99986	0.00014

$\hat{z} = \hat{a} + \hat{b}\xi$ $\hat{a} = +1.5336$ $\hat{b} = -1.1730$

D_n (tabulated) = 0.05862

D_n (calculated) = 0.04541

D_n (cal) < D_n (tab) .

Therefore, we can safely assume that the differences between the observed relative percentage fertility [$F_m(x)$] and the fitted one [$\hat{F}_m(x)$] is not significantly large.

References

1. Avery, R. C. (1970) "Graduation of Age-Specific Fertility Rates" Third Conference on the Mathematics of Population July 19-24, 1970, The University of Chicago.

2. Elderdon, E. P. (1953) "Frequency Curves and Correlation" Harren Press, Washington, D.C.

3. Hashmi, S. S. (1965) "The People of Karachi Demographic Characteristics" Monographs in Economics of Development No. 13, Pakistan Institute of Development of Economics, Karachi

4. Messay (1951) "The Kolmogrove-Simernove Test for Goodness-of-fit" Journal of American Statistical Association 46, pp. 70.

5. Mitra, S. (1967) "The Pattern of Age-Specific Fertility Rates" Demography 4(2):894-906.

6. Mitra, S. (1970) "Graduation of Net Maternity Function" Paper presented at the 1970 Meeting of the Population Association of America, Atlanta.

7. Pearson, E. S. Hartley, H. O. (1966) "Biometrika Tables for Statisticians" Cambridge University Press.

8. Pressat, R. (1972) "Demographic Analysis, Methods, Results and Applications Aldine-Atherton, New York.